

PF-0621 USN

<110> INCYTE GENOMIC, INC.; TANG, Y. Tom;  
CORLEY, Neil C.; GORGONE, Gina A.;  
GUEGLER, Karl J.; BAUGHN, Mariah R.

<120> MYOSIN HEAVY CHAIN HOMOLOG

<130> PF-0621 USN

<140> US 09/830,914

<141> To Be Assigned

<150> PCT/US99/26177

<151> 1999-11-05

<150> US 60/172,248

<151> 1998-11-05

<160> 4

<170> PERL Program

<210> 1

<211> 612

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 1929760CD1

<400> 1

Met	Phe	Cys	Pro	Pro	Gln	Val	Ser	Cys	Ser	Leu	Ser	Leu	Met	Pro
1				5					10					15
Arg	Leu	Pro	Ser	Ile	Arg	His	Trp	Gln	Gly	Pro	Ser	His	Pro	Gly
				20					25					30
Phe	Leu	Gly	Pro	Leu	Phe	Pro	Ile	Cys	Ser	Leu	Gln	Trp	Pro	His
				35					40					45
Gly	Phe	Ser	Ala	Ile	Phe	Pro	Gly	Leu	Leu	Asp	Val	Tyr	Gly	Phe
				50					55					60
Glu	Ser	Phe	Pro	Asp	Asn	Ser	Leu	Glu	Gln	Leu	Cys	Ile	Asn	Tyr
				65					70					75
Ala	Asn	Glu	Lys	Leu	Gln	Gln	His	Phe	Val	Ala	His	Tyr	Leu	Arg
				80					85					90
Ala	Gln	Gln	Glu	Glu	Tyr	Ala	Val	Glu	Gly	Leu	Glu	Trp	Ser	Phe
				95					100					105
Ile	Asn	Tyr	Gln	Asp	Asn	Gln	Pro	Cys	Leu	Asp	Leu	Ile	Glu	Gly
				110					115					120
Ser	Pro	Ile	Ser	Ile	Cys	Ser	Leu	Ile	Asn	Glu	Glu	Cys	Arg	Leu
				125					130					135
Asn	Arg	Pro	Ser	Ser	Ala	Arg	Gln	Leu	Gln	Thr	Arg	Ile	Glu	Thr
				140					145					150
Ala	Leu	Ala	Gly	Ser	Pro	Cys	Leu	Gly	His	Asn	Lys	Leu	Ser	Arg
				155					160					165
Glu	Pro	Ser	Phe	Ile	Val	Val	His	Tyr	Ala	Gly	Pro	Val	Arg	Tyr
				170					175					180

PF-0621 USN

His Thr Ala Gly	Leu Val Glu Lys Asn	Lys Asp Pro Ile Pro	Pro
	185	190	195
Glu Leu Thr Arg	Leu Leu Gln Gln Ser	Gln Asp Pro Leu Leu	Met
	200	205	210
Gly Leu Phe Pro	Thr Asn Pro Lys Glu	Lys Thr Gln Glu Glu	Pro
	215	220	225
Pro Gly Gln Ser	Arg Ala Pro Val Leu	Thr Val Val Ser Lys	Phe
	230	235	240
Lys Ala Ser Leu	Glu Gln Leu Leu Gln	Val Leu His Ser Thr	Thr
	245	250	255
Pro His Tyr Ile	Arg Cys Ile Lys Pro	Asn Ser Gln Gly Gln	Ala
	260	265	270
Gln Thr Phe Leu	Gln Glu Glu Val Leu	Ser Gln Leu Glu Ala	Cys
	275	280	285
Gly Leu Val Glu	Thr Ile His Ile Ser	Ala Ala Gly Phe Pro	Ile
	290	295	300
Arg Val Ser His	Arg Asn Phe Val Glu	Arg Tyr Lys Leu Leu	Arg
	305	310	315
Arg Leu His Pro	Cys Thr Ser Ser Gly	Pro Asp Ser Pro Tyr	Pro
	320	325	330
Ala Lys Gly Leu	Pro Glu Trp Cys Pro	His Ser Glu Glu Ala	Thr
	335	340	345
Leu Glu Pro Leu	Ile Gln Asp Ile Leu	His Thr Leu Pro Val	Leu
	350	355	360
Thr Gln Ala Ala	Ala Ile Thr Gly Asp	Ser Ala Glu Ala Met	Pro
	365	370	375
Ala Pro Met His	Cys Gly Arg Thr Lys	Val Phe Met Thr Asp	Ser
	380	385	390
Met Leu Glu Leu	Leu Glu Cys Gly Arg	Ala Arg Val Leu Glu	Gln
	395	400	405
Cys Ala Arg Cys	Ile Gln Gly Gly Trp	Arg Arg His Arg His	Arg
	410	415	420
Glu Gln Glu Arg	Gln Trp Arg Ala Val	Met Leu Ile Gln Ala	Ala
	425	430	435
Ile Arg Ser Trp	Leu Thr Arg Lys His	Ile Gln Arg Leu His	Ala
	440	445	450
Ala Ala Thr Val	Ile Lys Arg Ala Trp	Gln Lys Trp Arg Ile	Arg
	455	460	465
Met Ala Cys Leu	Ala Ala Lys Glu Leu	Asp Gly Val Glu Glu	Lys
	470	475	480
His Phe Ser Gln	Ala Pro Cys Ser Leu	Ser Thr Ser Pro Leu	Gln
	485	490	495
Thr Arg Leu Leu	Glu Ala Ile Ile Arg	Leu Trp Pro Leu Gly	Leu
	500	505	510
Val Leu Ala Asn	Thr Ala Met Gly Val	Gly Ser Phe Gln Arg	Lys
	515	520	525
Leu Val Val Trp	Ala Cys Leu Gln Leu	Pro Arg Gly Ser Pro	Ser
	530	535	540
Ser Tyr Thr Val	Gln Thr Ala Gln Asp	Gln Ala Gly Val Thr	Ser
	545	550	555
Ile Arg Ala Leu	Pro Gln Gly Ser Ile	Lys Phe His Cys Arg	Lys
	560	565	570
Ser Pro Leu Arg	Tyr Ala Asp Ile Cys	Pro Glu Pro Ser Pro	Tyr
	575	580	585

PF-0621 USN

```
Ser Ile Thr Gly Phe Asn Gln Ile Leu Leu Glu Arg His Arg Leu
                    590                    595                    600
Ile His Val Thr Ser Ser Ala Phe Thr Gly Leu Gly
                    605                    610
```

<210> 2

<211> 2109

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 1929760CB1

<400> 2

```
tgatgctctg ggctgtcttc acacttcatt tgggtttcct gcttgctctg agctctacag 60
gggaatgggg tagagatggg agccaccttg ggtggagggg ggggaaggta tgttctgccc 120
accacagggtg tcatgctcac tcagcctgat gcccaggctg ccaagtataa ggcatgggca 180
ggggcccagc caccctgggt tccttgggtcc cctattcccc atctgctccc tgcagtggcc 240
ccatgggttc tctgccatct tcccaggcct gctggatgtg tatggatttg aatcatttcc 300
tgacaacagt ctggaacagt tgtgcatcaa ctacgccaat gagaagctgc agcagcattt 360
tgtggctcac tacctaaggg cccagcagga ggaatacgca gttgagggcc tggagtggtc 420
attcatcaac taccaggaca accagccctg tttggatctc attgagggaa gccccatcag 480
catctgctcc ctcataaatg aggaatgccg cctcaatcga cccagcagcg cacgccagct 540
ccagacacgc attgagactg ccctggcagg cagcccctgc ctggggccaca ataagctcag 600
ccgggagccc agcttcattg tgggtgcatta tgcggggcct gtgcgggtacc acacagcagg 660
cctggtggag aagaacaagg accctatccc acctgagctg accaggctcc tgcagcaatc 720
ccaggacccc ctgctcatgg ggctgtttcc tactaacccc aaagagaaga cccaggagga 780
accccctggc cagagcaggg cccctgtgtt gaccgtgggt tccaagttca aggcctcact 840
ggagcagctt ctgcaggctc tacacagcac cagccccac tacattcgct gcatcaagcc 900
caacagccag ggccaggcgc agacctttct ccaagaggag gtccctgagcc agctggaggc 960
ctgtggcctc gtggagacca tccatatcag tgctgctggc ttccccatcc gggctctctca 1020
ccgaaacttt gtagaacgat acaagttact aagaaggctt catccttgca catcctctgg 1080
ccccgacagc ccatatcctg ccaaagggct ccctgaatgg tgtccacaca gcgaggaagc 1140
cacgcttgaa cctctcatcc aggacattct ccacactctg ccggtcctaa ctgaggcagc 1200
agccataact ggtgactcgg ctgaggccat gccagcccc atgcactgtg gcaggaccaa 1260
ggtgttcatg actgactcta tgctggagct tctggaatgt gggcgtgccc ggggtgctgga 1320
gcagtgtgcc cgctgcatcc aggggtggctg gaggcgacac cggcaccgag agcaggagcg 1380
gcagtggcgg gccgtcatgc tcatccaggc agccattcgt tcctggttaa ctcgaaaca 1440
catccagagg ctgcatgcag ctgccacagt catcaagcgt gcatggcaga agtgagaat 1500
cagaatggcc tgccttgctg ctaaagagct ggatggtgtg gaagaaaaac acttctctca 1560
agctccctgt tccctgagca cctcgccgct gcagaccagg ctcttgagg caataatccg 1620
cctctggccc ctgggactgg tcctggccaa tacggctatg ggtgtaggca gctttcagag 1680
gaaattagtg gtctgggctt gcctccagct cccagggggc agccccagta gctacactgt 1740
ccagacagca caagaccagg ctggtgtcac gtccatccga gcgctgcctc agggatcgat 1800
aaagtttcac tgcagaaagt ctccactgcy gtatgctgac atctgccctg aaccttcacc 1860
ctacagcatt acaggcttta atcagattct gctggaaaga cacaggctga tccacgtgac 1920
ctcttctgcc ttcactgggc tggggtgatc cttggtgcct ttgtttccac aaggcctttt 1980
cctgccccct gccttgccaa agacatttaa tcagcacaca gctgccagac tattcccaca 2040
gtgctccaaa tgcacatgaa caacagtgcg ggctccagcc ttcgaccag agccccgtgc 2100
ccagtgcgt                                     2109
```

<210> 3

<211> 1839

PF-0621 USN

<212> PRT

<213> *Caenorhabditis elegans*

<220>

<221> misc\_feature

<223> GenBank ID No: g1279777

<400> 3

Met	Phe	Asn	Tyr	Ser	Lys	Ile	Phe	Gln	Ile	His	Arg	Ala	Cys	Ser
1				5					10					15
Pro	Asn	Arg	Lys	Lys	Ile	Gly	Ser	Ile	Gln	Tyr	Gly	Arg	Arg	Arg
			20						25					30
His	Ser	Trp	Gln	Gly	Pro	Val	Val	Pro	Ala	Ala	Lys	Leu	Gln	Val
			35						40					45
Leu	Ile	Lys	Gly	Val	Arg	Ile	Trp	His	Arg	His	Pro	Thr	Leu	Val
			50						55					60
Trp	Ile	Gly	Ala	Thr	Leu	Glu	Glu	Asp	Ile	Thr	Phe	Gln	Thr	Arg
			65						70					75
Asn	Val	Arg	Ile	Arg	Leu	Glu	Asp	Asp	Thr	Glu	Val	Glu	Tyr	Ala
			80						85					90
Ile	Lys	Ser	Leu	Asp	Gln	Leu	Pro	Phe	Leu	Arg	Asn	Pro	Ala	Phe
			95						100					105
Leu	Val	Gly	Lys	Asp	Asp	Leu	Thr	Leu	Leu	Ser	Tyr	Leu	His	Glu
			110						115					120
Pro	Ala	Val	Leu	His	Asn	Leu	Gln	Val	Arg	Phe	Val	Lys	Gly	Ser
			125						130					135
Ser	Ile	Tyr	Thr	Tyr	Cys	Gly	Ile	Val	Leu	Val	Ala	Ile	Asn	Pro
			140						145					150
Tyr	Ala	Asp	Cys	Ser	His	Ile	Tyr	Gly	Glu	Glu	Ile	Ile	Gln	Val
			155						160					165
Tyr	Arg	Gly	Ala	Gly	Lys	Ser	Ala	Arg	Glu	Met	Asp	Pro	His	Ile
			170						175					180
Phe	Ala	Val	Ala	Glu	Glu	Ala	His	Phe	Asp	Met	Gly	Ala	Phe	Gly
			185						190					195
Lys	Ser	Gln	Ser	Ile	Ile	Val	Ser	Gly	Glu	Ser	Gly	Ala	Gly	Lys
			200						205					210
Thr	Val	Ser	Ala	Lys	Phe	Val	Met	Arg	Tyr	Leu	Ala	Ser	Val	Ala
			215						220					225
Ala	Ser	Lys	Thr	Arg	Asn	Gly	Gly	Thr	Thr	Ser	Ile	Glu	Ala	Arg
			230						235					240
Val	Leu	Ala	Ser	Asn	Pro	Ile	Met	Glu	Ser	Ile	Gly	Asn	Ala	Lys
			245						250					255
Thr	Ile	Arg	Asn	Asp	Asn	Ser	Ser	Arg	Phe	Gly	Lys	Phe	Ile	Gln
			260						265					270
Ile	Asn	Phe	Cys	Glu	Arg	Gly	Arg	Arg	Ile	Val	Gly	Ala	Glu	Met
			275						280					285
Lys	Thr	Tyr	Leu	Leu	Glu	Lys	Ser	Arg	Leu	Val	Phe	Gln	Ala	Pro
			290						295					300
Gly	Glu	Arg	Asn	Tyr	His	Ile	Phe	Tyr	Gln	Leu	Cys	Ala	Ala	Arg
			305						310					315
Asn	His	Gln	Val	Leu	Lys	Asp	Leu	His	Leu	Gly	Pro	Cys	Glu	Ser
			320						325					330
Tyr	Ser	Tyr	Leu	Thr	Gln	Gly	Gly	Asp	Ser	Arg	Ile	Pro	Gly	Val
			335						340					345

PF-0621 USN

Asp Asp Lys Ala Asp Phe Glu Ala Leu Leu Lys Ala Leu Gln Leu	350	355	360
Leu Gly Phe Asp Glu Lys Gln Met Ser Asp Val Phe Arg Leu Leu	365	370	375
Ala Gly Leu Leu Leu Leu Gly Asn Val His Phe Glu Asn Gly Glu	380	385	390
Gly Ser Ser Ala Val Ser Ala Ser Ser Cys Gln Glu Ile Ser Arg	395	400	405
Leu Cys Arg Glu Phe Trp Lys Ile Ser Glu Ser Asp Leu Arg Ile	410	415	420
Trp Leu Thr Arg Arg Glu Ile Arg Ala Val Asn Glu Ile Val Thr	425	430	435
Lys Pro Leu Thr Lys Asn Glu Ala Val Arg Ser Arg Asp Ala Leu	440	445	450
Thr Lys Met Leu Tyr Ser His Leu Phe Gly Trp Leu Val Asp Lys	455	460	465
Ile Asn Glu Ala Leu Asn Glu Lys Asp Lys Leu Asp Gly Thr Asn	470	475	480
Gln Lys Lys Arg Pro Asp Arg Phe Ile Gly Val Leu Asp Ile Tyr	485	490	495
Gly Phe Glu Thr Phe Asp Val Asn Ser Phe Glu Gln Phe Ser Ile	500	505	510
Asn Tyr Ala Asn Glu Lys Leu Gln Gln Gln Phe Asn Gln His Val	515	520	525
Phe Lys Leu Glu Gln Glu Glu Tyr Ile Arg Glu Glu Ile Glu Trp	530	535	540
Val Arg Val Asp Phe His Asp Asn Gln Pro Ala Ile Asp Leu Ile	545	550	555
Glu Gly Pro Val Gly Met Ile Asn Leu Leu Asp Glu Gln Cys Lys	560	565	570
Arg Leu Asn Gly Ser Asp Ala Asp Trp Leu Ser Gln Leu Gln Asn	575	580	585
Ser Thr Glu Leu Lys Arg Asn Pro Gln Leu Ala Phe Pro Lys Val	590	595	600
Arg Ser Asn Asp Phe Ile Val Arg His Phe Ala Ala Asp Val Thr	605	610	615
Tyr Ser Thr Asp Gly Phe Val Glu Lys Asn Arg Asp Ala Ile Gly	620	625	630
Glu Gln Leu Leu Asp Val Val Val Ala Ser Lys Phe Pro Phe Ile	635	640	645
Arg Thr Val Ile Gly Ser Thr Ala Pro Thr Ser Val Ser Ser Ser	650	655	660
Ser Ser Ser Ser Thr Pro Gly Lys Arg Thr Ile Lys Lys Thr Val	665	670	675
Ala Ser Gln Phe Arg Asp Ser Leu Lys Glu Leu Met Ser Val Leu	680	685	690
Cys Ser Thr Arg Pro His Tyr Val Arg Cys Ile Lys Pro Asn Asp	695	700	705
Ser Lys Ile Ser Phe Asp Phe Glu Pro Lys Arg Ala Ile Gln Gln	710	715	720
Leu Arg Ala Cys Gly Val Leu Glu Thr Val Arg Ile Ser Ala Ala	725	730	735
Gly Phe Pro Ser Arg Tyr Pro Tyr Glu Glu Phe Ala Arg Arg Tyr	740	745	750

PF-0621 USN

Arg Val Ile Tyr Thr Lys Glu Ala Ala Leu Trp Arg Asp Lys Pro	755	760	765
Lys Gln Phe Ala Glu Leu Ala Cys Gln Gln Cys Leu Glu Glu Gly	770	775	780
Lys Tyr Ala Val Gly Lys Thr Lys Ile Phe Leu Arg Thr Gly Gln	785	790	795
Val Ala Val Leu Glu Arg Val Arg Leu Asp Thr Leu Ala Ala Ala	800	805	810
Ala Thr Val Ile Gln Lys Met Trp Lys Gly Phe Leu Ala Arg Arg	815	820	825
Lys Tyr Glu Thr Met Arg Arg Ser Leu Leu Ile Val Gln Ala Ser	830	835	840
Leu Lys Ala Phe Leu Ala Phe Arg Arg Ile Lys Tyr Leu Gln Met	845	850	855
His Arg Ala Val Ile Val Met Gln Ser Ala Val Arg Gly Tyr Leu	860	865	870
Glu Arg Arg Lys Tyr Glu Gln Ile Arg Asp Ser Ile Ile Gly Ile	875	880	885
Gln Ala Met Phe Lys Ala Asn Arg Val Arg Arg Tyr Val Glu Lys	890	895	900
Leu Arg Tyr Glu Lys Ser Ala Ile Thr Ile Gln Ala Ala Trp Arg	905	910	915
Gly Tyr Leu Ala Arg Arg Glu Gln Ile Ala Asn Arg Lys Lys Val	920	925	930
Val Met Val Gln Cys Ala Val Arg Lys Trp Leu Ala Lys Arg Arg	935	940	945
Leu Arg Glu Leu Lys Ile Glu Ala Arg Ser Val Gly His Leu Gln	950	955	960
Lys Leu Asn Thr Gly Leu Glu Asn Lys Ile Ile Glu Leu Gln Met	965	970	975
Arg Leu Asp Ile Ala Asn Ala Arg Thr Lys Glu Glu Ala Glu Lys	980	985	990
Phe Ala Thr Ala Ser Lys Asn Leu Gln Lys Thr Lys Ala Asp Leu	995	1000	1005
Ala Met Met Glu Ala Glu Arg Leu Thr Leu Leu Glu Ala Arg Asn	1010	1015	1020
Arg Val Glu Val Leu Gln Glu Glu Val Glu Arg Leu Glu Thr Glu	1025	1030	1035
Cys Asp Leu Lys Glu Ala Gln Arg Gly Gly Met Glu Thr Lys Met	1040	1045	1050
Val Glu Leu Gln Ser Arg Leu Asp Gln Phe Gln Met Gln Ser Glu	1055	1060	1065
Ser Gly Gln Thr Ile Val Glu Leu Thr Glu Gln Leu Glu Lys Ala	1070	1075	1080
Lys Ala Asp Arg Val Leu Trp Asp Glu Glu Arg Gln Arg Met Glu	1085	1090	1095
Ala Ala Leu Asn Thr Glu Arg Ser Ala Arg Asn Ala Leu Asp Ala	1100	1105	1110
Glu Met Ala Ala Met Arg Glu Gln Leu Met Lys Asn Val Asp Leu	1115	1120	1125
Phe Glu Ser Ser Thr Phe Gln Lys Arg Pro Ser Gln Lys Lys Asn	1130	1135	1140
Arg Asp Asp Asp Ser Cys Ser Arg Thr Thr Ser Asn Leu Ser Gln	1145	1150	1155

PF-0621 USN

Leu Thr Gly Ser Phe Thr Ala Glu Thr Ile Asn Gly Val His Ser	1160	1165	1170
Thr Ser Arg Gly Ser Pro Glu Val Leu Leu Asp Asn Met Ala Ser	1175	1180	1185
Thr Phe Glu Gln Leu Arg Met Ile Asn Asp Leu Arg Gln Arg Asn	1190	1195	1200
Glu His Cys Gln Arg Glu Thr Glu Arg Met Lys Ala Ile Ile Glu	1205	1210	1215
Ala Ser Thr Leu Ile Glu Thr Leu Asp Lys Lys Thr Ser Leu Lys	1220	1225	1230
Ala Phe Glu Ser Ile Arg Val Gly Glu Leu Glu Gly Ala Tyr Asn	1235	1240	1245
Arg Leu Lys Asn Asp Met Glu Arg Leu Val Ser Gly Glu Asn Gly	1250	1255	1260
Ala Thr His Ser Val Phe Glu Arg Ile Met Glu Glu Asn Glu Arg	1265	1270	1275
Leu Arg Glu Glu Ala Val Glu Leu Arg Ser Met Leu Ser Ser His	1280	1285	1290
Phe Glu Lys Gln Ser Val Ala Gly Ser Ser Gly Tyr Arg Arg Ser	1295	1300	1305
Pro Arg Pro Asp Ser Gly His Cys Ser Gly Ala Asp Ser Glu Asp	1310	1315	1320
Gly Ser Ser Gly Ala Asp Leu Glu Glu Asp Leu Cys Ile Glu Arg	1325	1330	1335
Gln Cys Arg His Leu Lys Asn Leu Ala Glu Asn Leu Thr Lys Met	1340	1345	1350
Leu Thr Asn Gln Asn Leu Glu Ile Glu Arg Leu Gln Gln Gln Leu	1355	1360	1365
Arg Phe Ser Glu Ser Gln Thr Val Phe Arg Pro Ser Asp Cys Ser	1370	1375	1380
Leu Asp Glu Ala Val Arg Gly Ala His Lys Gln Thr Gln Leu Leu	1385	1390	1395
Ala Gln Gln Asn Met Asp Leu Asn Asp Lys Leu Thr Arg Gln Ser	1400	1405	1410
Glu Glu Leu Ala Glu Ala Arg Ala Gln Leu Arg Gly Tyr Ser Gly	1415	1420	1425
Pro Leu Gly Leu Glu Asn Ala Ser Asp Glu Glu Ile Ile Arg Leu	1430	1435	1440
Glu Ala Phe Glu Lys Gly Ser Ile Lys His Ser Gly Phe Leu Glu	1445	1450	1455
Val Tyr Asn Val Pro Glu Phe Ala Arg Ile Ile Val Cys Glu Leu	1460	1465	1470
Lys Pro Thr Leu Ala Arg Leu Leu Thr Lys Asn Leu Pro Ala Tyr	1475	1480	1485
Leu Leu Val Ala Ala Phe Arg Asn His Asp Glu Lys Arg Asp Glu	1490	1495	1500
Thr Ala Leu Thr Gly Leu Phe Ser Ser Val His Leu Val Leu Lys	1505	1510	1515
Asp Thr Ile Ser Arg Ser His Asp Leu Asp Leu Leu Ser Leu Trp	1520	1525	1530
Leu Val Asn Leu Trp Arg Leu Phe Asn Leu Leu Arg Gln Tyr Ser	1535	1540	1545
Gly Glu Asp Ser Gln Pro Glu Trp His Val Ala Asn Thr Glu Thr	1550	1555	1560

Gln	Asn	Ser	Tyr	Arg	Phe	Lys	Ala	Tyr	Asp	Val	Ala	Pro	Ile	Arg	
									1565	1570					1575
Asp	Gln	Leu	Lys	Leu	Arg	Ile	Glu	Glu	Cys	Tyr	Thr	Ser	Leu	Met	
									1580	1585					1590
Lys	Lys	Ala	Ile	Glu	His	Val	Leu	Ser	Pro	Lys	Ile	Val	Pro	Gly	
									1595	1600					1605
Ile	Leu	Gln	His	Glu	Ser	Ser	Ser	Asp	Leu	Met	Thr	Ala	Gly	Gln	
									1610	1615					1620
Glu	Arg	Arg	Asp	Arg	Asn	Ser	Gly	Ser	Val	Glu	Ser	Gln	Arg	Lys	
									1625	1630					1635
Ser	Leu	Asp	Asp	Leu	Leu	Gln	Phe	Met	Glu	Ile	Val	His	Thr	Lys	
									1640	1645					1650
Leu	Thr	Thr	Tyr	Gly	Gly	Asp	Asp	Ile	Val	Val	Lys	Gln	Val	Ile	
									1655	1660					1665
Gly	Gln	Met	Ala	Arg	Trp	Met	Cys	Ala	Leu	Ala	Leu	Asn	Tyr	Met	
									1670	1675					1680
Met	Phe	Arg	Arg	Glu	Leu	Cys	Asn	Phe	Glu	Lys	Ala	Ile	Gln	Ile	
									1685	1690					1695
Lys	His	Asn	Val	Thr	Gln	Ile	Gln	Asn	Trp	Leu	Asn	Ala	Lys	Gly	
									1700	1705					1710
Leu	Ser	Asp	Cys	Arg	Asp	His	Phe	Glu	Pro	Leu	Val	Gln	Ala	Cys	
									1715	1720					1725
His	Leu	Leu	Gln	Ser	Arg	Lys	Asp	Pro	Ser	Asn	Leu	Asp	Thr	Leu	
									1730	1735					1740
Cys	Gly	Glu	Met	Thr	Ser	Arg	Leu	Lys	Pro	Arg	Gln	Val	Val	Ala	
									1745	1750					1755
Ile	Leu	Gln	His	Tyr	Asp	Pro	Ser	Asp	Glu	Met	Glu	Asp	Gly	Leu	
									1760	1765					1770
Ser	Pro	Glu	Phe	Leu	Val	Gln	Ile	Gln	Lys	Lys	Leu	Asn	Glu	Arg	
									1775	1780					1785
Ala	Ile	Ala	Asn	Asn	Asp	Pro	Ile	Glu	Asp	Lys	Asp	Lys	Leu	Ile	
									1790	1795					1800
Met	Leu	Gly	Thr	Tyr	Leu	Pro	Pro	Phe	Asp	Thr	Gln	Pro	Phe	Ser	
									1805	1810					1815
Tyr	Ser	Asp	Phe	Pro	Leu	Glu	Thr	Leu	Ser	Leu	Pro	Ser	Cys	Leu	
									1820	1825					1830
His	Met	Gln	Ser	Val	Cys	Arg	Leu	Val							
									1835						

$\langle 211 \rangle$  1120

&lt;212&gt; PRT

<213> Helianthus annuus

<220>

```
<221> misc_feature
```

<223> GenBank ID No: g2444174

<400> 4

Met	Asp	Arg	Val	Val	Asp	Asp	Asp	Ser	Pro	Tyr	Gly	Gln	Gly	Ser
1				5					10					15
Ser	Phe	Leu	Leu	Asn	Asp	Arg	Pro	Ser	Val	Asp	Asp	Val	Asn	Asp
				20					25					30
Asp	Asp	Asp	Ala	Asp	Val	Asn	Pro	Ser	Val	Ser	Ala	Gln	Gly	Ser



	35	40	45
Val Leu Gly Ser Trp Gly Asn Lys Lys Trp Gly Asp Thr Ala Ser			
	50	55	60
Tyr Ile Ala Lys Lys Lys Leu Gln Ser Trp Phe Gln Thr Ser Asp			
	65	70	75
Gly Asn Trp Glu Leu Ala Lys Ile Leu Ser Ile Thr Gly Ser Glu			
	80	85	90
Ser Leu Met Ser Leu Ser Glu Glu Lys Val Leu Lys Val Ser Ser			
	95	100	105
Asp Ser Leu Leu Pro Ala Asn Pro Glu Ile Leu Asp Gly Val Asp			
	110	115	120
Asp Leu Met Gln Leu Ser Tyr Leu Asn Glu Pro Ser Val Leu Tyr			
	125	130	135
Asn Leu Gln Tyr Arg Tyr Asp Arg Asp Met Ile Tyr Ser Lys Ala			
	140	145	150
Gly Pro Val Leu Val Ala Ile Asn Pro Phe Lys Lys Ile Pro Leu			
	155	160	165
Tyr Gly Ser Asp Tyr Ile Glu Ala Tyr Lys Arg Lys Ser Ile Asp			
	170	175	180
Asn Pro His Val Tyr Ala Ile Ala Asp Thr Ala Ile Arg Glu Met			
	185	190	195
Ile Arg Asp Glu Val Asn Gln Ser Ile Val Ile Ser Gly Glu Ser			
	200	205	210
Gly Ala Gly Lys Thr Glu Thr Pro Lys Ile Ala Met Gln Tyr Leu			
	215	220	225
Ala Ala Leu Gly Gly Gly Asp Ala Arg Glu Ser Gly Ile Leu Ser			
	230	235	240
His Asn Gly Cys Arg Thr Pro Arg Arg Ala Glu Ala Phe Gly Asn			
	245	250	255
Ala Lys Thr Ser Arg Asp Asn Asn Ser Ser Arg Ile Gly Lys Leu			
	260	265	270
Ile Glu Ile His Phe Ser Glu Thr Gly Lys Ile Ser Gly Ala Lys			
	275	280	285
Ile Gln Thr Phe Leu Leu Glu Lys Ser Arg Val Val Gln Cys Thr			
	290	295	300
Asp Gly Glu Arg Ser Tyr His Ser Phe Tyr Gln Leu Cys Ala Gly			
	305	310	315
Ala Pro Pro Ser Leu Arg Glu Lys Leu Asn Leu Lys Ser Ala Arg			
	320	325	330
Glu Tyr Lys Tyr Phe Gln Gln Ser Thr Cys Tyr Ser Ile Asn Gly			
	335	340	345
Val Asp Asp Ala Glu Glu Phe Arg Val Val Val Glu Ala Leu Asp			
	350	355	360
Ala Val His Val Ser Lys Glu Asn Gln Glu Asn Ala Phe Ala Met			
	365	370	375
Leu Ala Ala Val Leu Trp Leu Gly Asn Val Thr Phe Ser Ile Val			
	380	385	390
Asp Asn Glu Asn His Val Glu Pro Ile Ile Asp Asp Ala Leu Leu			
	395	400	405
Asn Val Ala Lys Leu Ile Gly Cys Glu Ala Asp Asp Leu Lys Leu			
	410	415	420
Ala Leu Ser Thr Arg Asn Met Lys Val Gly Asn Asp Ile Ile Val			
	425	430	435
Gln Lys Leu Thr Leu Ala Gln Ala Ile Asp Thr Arg Asp Ala Leu			

Ala Lys Ser Ile Tyr Ser Cys Leu Phe	440	445	450
Asp Trp Leu Val Glu Gln	455	460	465
Ile Asn Lys Ser Leu Ala Val Gly Lys Arg Arg Thr Gly Arg Ser	470	475	480
Ile Ser Ile Leu Asp Ile Tyr Gly Phe Glu Ser Phe Asp Val Asn	485	490	495
Ser Phe Glu Gln Phe Cys Ile Asn Tyr Ala Asn Glu Arg Leu Gln	500	505	510
Gln His Phe Asn Arg His Leu Phe Lys Leu Glu Gln Glu Glu Tyr	515	520	525
Ile Gln Asp Gly Ile Asp Trp Ala Lys Val Asp Phe Glu Asp Asn	530	535	540
Gln Asp Cys Leu Asn Leu Phe Glu Lys Lys Pro Leu Gly Leu Met	545	550	555
Thr Leu Leu Asp Glu Glu Ser Thr Phe Pro Asn Gly Thr Asp Met	560	565	570
Thr Phe Ala Thr Lys Leu Lys Gln His Leu Lys Thr Asn Ser Cys	575	580	585
Phe Arg Gly Glu Arg Gly Lys Ala Phe Thr Val His His Tyr Ser	590	595	600
Gly Glu Val Thr Tyr Asp Thr Ser Gly Phe Leu Glu Lys Asn Arg	605	610	615
Asp Leu Leu His Leu Asp Ser Ile Gln Leu Leu Ser Ser Cys Thr	620	625	630
Cys Glu Leu Pro Gln Ala Phe Ala Ser Asn Met Leu Ser Leu Ser	635	640	645
Glu Lys Pro Val Pro Gly Pro Leu His Lys Ser Gly Gly Ala Asp	650	655	660
Ser Gln Lys Leu Ser Val Val Thr Lys Phe Lys Gly Gln Leu Phe	665	670	675
Gln Leu Met Gln Arg Leu Glu Ser Thr Thr Pro His Phe Ile Arg	680	685	690
Cys Ile Lys Pro Asn Asn Ser Gln Ser Pro Gly Ile Tyr His Gln	695	700	705
Gly Leu Val Leu Gln Gln Leu Arg Cys Cys Gly Val Leu Glu Val	710	715	720
Val Arg Ile Ser Arg Ser Gly Phe Pro Thr Arg Met Ser His Gln	725	730	735
Lys Phe Ala Arg Arg Tyr Gly Phe Leu Leu Leu Glu His Val Ala	740	745	750
Ser Gln Asp Pro Leu Ser Val Ser Val Ala Ile Leu His Gln Phe	755	760	765
Asp Ile Leu Pro Glu Met Tyr Gln Ile Gly Tyr Thr Lys Leu Phe	770	775	780
Phe Arg Thr Gly Gln Ile Gly Lys Leu Glu Asp Thr Arg Asn Arg	785	790	795
Thr Leu Asn Gly Ile Leu Arg Val Gln Ser Cys Phe Arg Gly His	800	805	810
Lys Ala Arg Gln Tyr Met Lys Glu Leu Lys Arg Gly Ile Phe Asn	815	820	825
Leu Gln Ala Phe Ala Arg Gly Glu Lys Thr Arg Lys Glu Phe Ala	830	835	840
Ile Leu Val His Arg His Arg Ala Ala Val His Ile Gln Lys His			

PF-0621 USN

	845		850		855
Ile Lys Ala Lys	Ile Ser Lys Lys Arg	Phe Glu Asp Val His	Gly		
	860		865		870
Ala Thr Ile Thr	Leu Gln Ala Val Ile	Arg Gly Trp Leu Val	Arg		
	875		880		885
Arg Cys Ser Gly	Asp Ile Ala Leu Leu	Gln Phe Gly Ser Gly	Lys		
	890		895		900
Gly Asn Gly Ser	Asp Glu Val Leu Val	Lys Ser Ser Tyr Leu	Ala		
	905		910		915
Glu Leu Gln Arg	Arg Ile Leu Lys Ala	Glu Ala Gly Leu Arg	Glu		
	920		925		930
Lys Glu Glu Glu	Asn Asp Ile Leu His	Gln Arg Leu Gln Gln	Tyr		
	935		940		945
Glu Asn Arg Trp	Ser Glu Tyr Glu Leu	Lys Met Lys Ser Met	Glu		
	950		955		960
Glu Val Trp Gln	Lys Gln Met Arg Ser	Leu Gln Ser Ser Leu	Ser		
	965		970		975
Ile Ala Lys Lys	Ser Leu Ser Tyr Asp	Ser Glu Arg Asn Ser			
	980		985		990
Asp Ala Ser Ile	Asn Thr Ala Asn Asp	Glu Thr Asn Pro Pro	Trp		
	995		1000		1005
Asp Ala Ala Thr	Asn Gly Arg Arg Asn	Gly Val Glu Asn Val	Arg		
	1010		1015		1020
Pro Met Ser Ala	Gly Leu Ser Val Ile	Ser Arg Leu Ala Glu	Glu		
	1025		1030		1035
Phe Glu Gln Arg	Ser Gln Val Phe Gly	Asp Asp Ala Lys Phe	Leu		
	1040		1045		1050
Val Glu Val Lys	Ser Gly Gln Val Glu	Ala Asn Leu Asn Pro	Asp		
	1055		1060		1065
His Glu Leu Arg	Arg Leu Lys Gln Met	Phe Glu Gly Trp Lys	Lys		
	1070		1075		1080
Asp Tyr Thr Ala	Arg Leu Arg Glu Thr	Lys Val Ile Leu Asn	Lys		
	1085		1090		1095
Leu Gly His Glu	Asp Gly Asp Gly Glu	Lys Gly Lys Lys Lys	Trp		
	1100		1105		1110
Trp Gly Arg Leu	Asn Ser Ser Arg Val	Asn			
	1115		1120		